



Multi-State Standard Setting Report

Praxis World Languages: German (0183)

Praxis World Languages: French (0174)

Praxis World Languages: Spanish (0195)

October 2009

Conducted by
Educational Testing Service
Princeton, New Jersey

Executive Summary

To support the decision-making process for state departments of education with regards to establishing passing scores, or cut scores, for the Praxis World Languages: German, French and Spanish assessments, research staff from Educational Testing Service (ETS) designed and conducted a series of multi-state standard setting studies. The studies also collected content-related validity evidence to confirm the importance of the content specifications for entry-level K-12 German, French and Spanish teachers.

Recommended Cut Scores

The standard setting studies involved two expert panels for each assessment, comprised of teachers, administrators and college faculty. The recommended cut scores for each panel, as well as the average cut score across the two panels, are provided to help state departments of education determine appropriate cut (or passing) scores.

- For Praxis World Languages: **German**, the average recommended cut score is **64** (on the raw score metric), which represents 65% of total available 98 raw points (the recommended cut scores for Panels 1 and 2 are 66 and 63, respectively). The scaled score associated with a raw score of 64 on the Praxis German assessment is 163.
- For Praxis World Languages: **French**, the average recommended cut score is **63** (on the raw score metric), which represents 65% of total available 97 raw points (the recommended cut scores for Panels 1 and 2 are 59 and 66, respectively). The scaled score associated with a raw score of 63 on the Praxis French assessment is 162.
- For Praxis World Languages: **Spanish**, the recommended cut score is **67** (on the raw score metric), which represents 70% of total available 96 raw points (the recommended cut scores for Panels 1 and 2 are 66 and 69, respectively). The scaled score associated with a raw score of 67 on the Praxis Spanish assessment is 168.

Summary of Specification Judgments

Panelists judged the extent to which the knowledge and/or skills reflected by the Praxis World Languages assessments content specifications were important for entry-level World Language teachers. For each assessment, all the knowledge/skills statements comprising the test specifications were judged to be *Very Important* or *Important* by a majority of the panelists, providing additional evidence that the content of the Praxis World Languages assessments is important for beginning practice.

Introduction

To support the decision-making process for state departments of education with regards to establishing passing scores, or cut scores, for the Praxis World Languages: German, French and Spanish assessments, research staff from Educational Testing Service (ETS) designed and conducted a series of multi-state standard setting studies. The studies also collected content-related validity evidence to confirm the importance of the content specifications for entry-level K-12 German, French and Spanish teachers. The standard setting studies involved two expert panels for each assessment, comprised of teachers, administrators, and college faculty. Panelists were recommended by departments of education of states that (a) currently use the Praxis Content Knowledge and/or Productive Language Skills assessments or (b) are considering use of the new Praxis World Languages assessments as part of their licensure process.

The design of the multi-state standard setting studies included two, non-overlapping panels to (a) allow each participating state to be represented and (b) replicate the judgment process to strengthen the technical quality of the recommended passing score for each assessment. (See Appendix A for the common agenda used for all panels.)

- **German:** Two non-overlapping panels with a total of 32 panelists representing 16 states (see Figure 1a) participated.
- **French:** Two non-overlapping panels with a total of 47 panelists representing 22 states (see Figure 1b) participated.
- **Spanish:** Two non-overlapping panels with a total of 39 panelists representing 23 states (see Figure 1c) participated.

Figure 1a. Participating States (and number of panelists) for German

Alabama (1 panelist)	South Carolina (2 panelists)
Delaware (1 panelist)	South Dakota (4 panelists)
Kentucky (2 panelists)	Tennessee (4 panelists)
Maryland (1 panelist)	Utah (2 panelists)
Mississippi (2 panelists)	Wisconsin (1 panelist)
North Carolina (2 panelists)	West Virginia (2 panelists)
North Dakota (4 panelists)	Wyoming (1 panelist)
Pennsylvania (2 panelists)	Nevada (1 panelist)

NOTE: Alabama, Delaware, Kentucky, Maryland, Wisconsin, Wyoming, and Nevada were represented on only one of the two panels.

Figure 1b. Participating States (and number of panelists) for French

Connecticut (2 panelists)	North Dakota (2 panelists)
Hawaii (1 panelist)	Pennsylvania (4 panelists)
Kentucky (4 panelists)	Rhode Island (1 panelist)
Louisiana (3 panelists)	South Carolina (3 panelists)
Maine (1 panelist)	South Dakota (1 panelist)
Maryland (3 panelists)	Tennessee (3 panelists)
Mississippi (4 panelists)	Utah (2 panelists)
Missouri (1 panelist)	Vermont (2 panelists)
Nevada (2 panelists)	Washington, D.C. (1 panelist)
New Hampshire (1 panelist)	West Virginia (2 panelists)
North Carolina (2 panelists)	Wisconsin (2 panelists)

NOTE: Hawaii, Maine, Missouri, New Hampshire, Rhode Island, South Dakota, Vermont, and Washington, D.C., were represented on only one of the two panels.

Figure 1c. Participating States (and number of panelists) for Spanish

Alabama (2 panelists)	North Dakota (2 panelists)
Delaware (1 panelist)	Ohio (1 panelist)
Hawaii (2 panelists)	Pennsylvania (2 panelists)
Kentucky (2 panelists)	South Carolina (2 panelists)
Louisiana (2 panelists)	South Dakota (2 panelists)
Maine (2 panelists)	Tennessee (1 panelist)
Maryland (2 panelists)	Utah (1 panelist)
Mississippi (2 panelists)	Vermont (3 panelists)
Missouri (1 panelist)	Washington, D.C. (1 panelist)
Nevada (1 panelist)	West Virginia (3 panelists)
New Hampshire (1 panelist)	Wisconsin (1 panelist)
North Carolina (2 panelists)	

NOTE: Delaware, Missouri, Nevada, New Hampshire, Ohio, Pennsylvania, Tennessee, Utah, Washington, D.C., and Wisconsin were represented on only one of the two panels.

The training provided to panelists as well as the study materials were consistent across panels within an assessment with the exception of defining the “just qualified candidate.” To assure that both panels for an assessment were using the same frame of reference when making question-level standard setting judgments, the “just qualified candidate” definition developed through a consensus process by the first panel was used as the definition for the second panel. The second panel did complete a thorough review of the definition to allow panelists to internalize the definition. The processes for developing the definition (with Panel 1) and reviewing/internalizing the definition (with Panel 2) are described later, and the “just qualified candidate” definitions are presented in Appendix B.

The panels were convened in July and August 2009 in Princeton, New Jersey. The results for each panel and results combined across panels for each assessment are summarized in the following report. The technical report containing the recommended passing scores for the German, French, and Spanish assessments is provided to each of the represented state departments of education. In each state, the department of education, the state board of education, or a designated educator licensure board is responsible for establishing the final passing scores in accordance with applicable state regulations.

The first national administration of the new Praxis World Languages assessments will occur in fall 2010. The current Praxis Content Knowledge and Productive Language Skills assessments will be phased out, with the last national administration in June 2010 for German and July 2010 for French and Spanish.

Praxis World Languages Assessments

The Praxis World Languages *Test at a Glance* documents (ETS, in press) for the German, French, and Spanish assessments describe the purpose and structure of the assessment. In brief, each assessment measures whether entry-level German, French, or Spanish teachers have the knowledge believed necessary for competent professional practice. A National Advisory Committee of expert practitioners and preparation faculty defined the content of the assessments, and a national survey of the field confirmed the content.

For each of the German, French, and Spanish assessments, the two hour and forty-five minute assessment is divided into four separately timed sections:

- **Section I: Listening with Cultural Knowledge** (50 minutes) – 36 multiple-choice questions¹
- **Section II: Reading with Cultural Knowledge** (50 minutes) – 39 multiple-choice questions².
- **Section III: Writing** (50 minutes) – Three constructed-response questions
- **Section IV: Speaking** (15 minutes) – Three constructed-response questions.

Candidate scores on the four sections are combined and reported as an overall score; five category scores – Listening, Reading, Cultural Knowledge, Writing, and Speaking – also are reported. The maximum total number of raw points that may be earned on each assessment is 98 for German, 97 for French, and 96 for Spanish. The reporting scales for the Praxis German, French, and Spanish assessments range from 100 to 200 scaled-score points.

¹ For Section I (Listening), 30 of the 36 questions contribute to the candidate's score.

² For Section II (Reading), 32 of the 39 questions contribute to the candidate's score for German; 31 of the 39 questions for French; and 30 of the 39 questions for Spanish.

Expert Panels

For each Praxis World Languages assessment, the standard setting study included two expert panels. The various state departments of education recruited panelists to represent a range of professional perspectives. A description of the panels for each assessment is presented below. (See Appendix C for a listing of panelists for each of the six panels.)

Praxis German Assessment

Panel 1 included 15 teachers, administrators, and college faculty who prepare K-12 German teachers, representing 11 states. In brief, 14 panelists were teachers and one was college faculty. Thirteen panelists were female. Nine panelists indicated they were most fluent in English, and five indicated they were equally fluent in English and German. Fourteen panelists reported being certified German teachers in their states. Approximately half of the panelists had between 4 and 11 years of experience as a K-12 German teacher, and approximately a quarter had 16 or more years of teaching experience.

Panel 2 included 17 teachers, administrators, and college faculty, representing 14 states. In brief, 14 panelists were teachers, one was an administrator, and two were college faculty. Twelve panelists were female. Twelve panelists indicated they were most fluent in English, and five indicated they were equally fluent in English and German. Approximately half of the panelists had 12 or more of experience as a K-12 German teacher, and approximately 20 percent had 3 or fewer years of teaching experience.

A fuller demographic description for the members of the two German panels is presented in Table 1 in Appendix D.

Praxis French Assessment

Panel 1 included 23 teachers, administrators, and college faculty who prepare K-12 French teachers, representing 18 states. In brief, 15 panelists were teachers, two were administrators, and five were college faculty. Nineteen panelists were White, three were African American, and one was Alaskan Native/American Indian. Seventeen panelists were female. Fourteen panelists indicated they were most fluent in English, and seven indicated they were equally fluent in English and French. Nineteen panelists reported being certified French teachers in their states. Approximately half of the panelists had between 4 and 11 years of experience as a K-12 French teacher, and over a third had 16 or more years of teaching experience.

Panel 2 included 24 teachers, administrators, and college faculty, representing 18 states. In brief, 19 panelists were teachers, two were administrators, and two were college faculty. Nineteen panelists were White, three were African American, and one was Asian American. Eighteen panelists were female. Nineteen panelists indicated they were most fluent in English, and two indicated they were equally fluent in English and French.

Approximately half of the panelists had 16 or more of experience as a K-12 French teacher, and approximately a quarter had 7 or fewer years of teaching experience.

A fuller demographic description for the members of the two French panels is presented in Table 7 in Appendix E.

Praxis Spanish Assessment

Panel 1 included 18 teachers, administrators, and college faculty who prepare K-12 Spanish teachers, representing 17 states. In brief, 12 panelists were teachers, two were administrators, and four were college faculty. Nine panelists were White, five were Hispanic, three were African American, and one was Asian American. Twelve panelists were female. Thirteen panelists indicated they were most fluent in English, and four indicated they were equally fluent in English and Spanish. Fourteen panelists reported being certified Spanish teachers in their states. Half of the panelists had 16 or more years of experience as a K-12 Spanish teacher, and nearly 40 percent had 11 or fewer years of teaching experience.

Panel 2 included 21 teachers, curriculum specialists, and college faculty, representing 19 states. In brief, 12 panelists were teachers, five were administrators, and four were college faculty. Eight panelists were White, eight were Hispanic, four were African American, and one was Native Hawaiian/Pacific Islander. Sixteen panelists were female. Ten panelists indicated they were most fluent in English, and nine indicated they were equally fluent in English and Spanish. Approximately half of the panelists had 16 or more of experience as a K-12 Spanish teacher, and more than 40 percent had 11 or fewer years of teaching experience.

A fuller demographic description for the members of the two Spanish panels is presented in Table 13 in Appendix F.

Process and Method

The design of the Praxis World Languages assessments standard setting studies included two non-overlapping expert panels for each assessment. As described below, the training provided to panelists and study materials were consistent across panels. Any differences between panels (e.g., defining the “just qualified candidate”) are highlighted.

The panelists were sent an e-mail explaining the purpose of the standard-setting study and requesting that they review the test content specifications for the Praxis World Languages assessment (included in the Praxis World Languages *Test at a Glance*, which was attached to the e-mail). The purpose of the review was to familiarize the panelists with the general structure and content of the assessment.

The standard-setting studies began with a welcome and introduction by Drs. Clyde Reese, Patricia Baron, and Wanda Swiggett, ETS researchers in the Center for Validity Research. Dr. Reese, lead facilitator for the studies, then explained how the particular Praxis World Language assessment was developed, provided an overview of standard setting, and presented the agenda for the study.

Reviewing the Praxis World Languages Assessments

The first activity was for the panelists to “take the test.” (Each panelist had signed a nondisclosure form.) The panelists were given approximately two hours to respond to the multiple-choice questions and to sketch responses to the constructed-response questions. The panelists had access to the answer key for the multiple-choice questions and access to the rubrics for the constructed response questions. The purpose of “taking the test” was for the panelists to become familiar with the test format, content, and difficulty.

The panelists then engaged in a discussion of the major content areas being addressed by the assessment; they were also asked to remark on any content areas that they thought would be particularly challenging for entering German, French, or Spanish teachers, and areas that addressed content that would be particularly important for entering teachers.

Defining the JQC

Following the review of the assessment, panelists internalized the definition of the Just Qualified Candidate (JQC). The JQC is the test taker who has the minimum level of skills believed necessary to be a qualified K-12 German, French, or Spanish teacher. The JQC definition is the operational definition of the cut score. The goal of the standard-setting process is to identify the test score that aligns with this definition of the JQC.

In Panel 1, the panelists were split into smaller groups, and each group was asked to write down their definition of a JQC. Each group referred to Praxis World Languages *Test at a Glance* to guide their definition. Each group posted its definition on chart paper, and a full-panel discussion occurred to reach consensus on a final definition (Appendix B).

In Panel 2, the panelists began with the definition of the JQC developed by the first panel. Given that each multi-state standard setting study was designed to replicate processes and procedures across the two panels for each assessment, it was important that both panels for an assessment use the same JQC definition to frame their judgments. For Panel 2, the panelists reviewed the JQC definition, and any ambiguities were discussed and clarified. The panelists then were split into smaller groups, and each group discussed the behaviors they would expect of the JQC based on the definition and developed performance indicators or “can do” statements based on the definition. The performance indicators were shared across groups and discussed. The purpose of the exercises was to have the panelists internalize the definition.

Panelists' Judgments

The standard-setting process for the Praxis World Languages assessments was conducted for the overall test, though one standard-setting approach was implemented for Sections I and II (multiple-choice questions) and another approach was implemented for Sections III and IV (constructed-response questions). Each panel's passing score for the assessment is the sum of the interim cut scores recommended by the panelists for each section. These approaches are described next, followed by the results from each standard-setting study. The recommended cut scores for each panel, as well as the average cut score across the two panels, are provided to help state departments of education determine appropriate cut (or passing) scores.

Standard Setting for Sections I and II (Multiple-Choice Questions). A probability-based Angoff method (Brandon, 2004; Hambleton & Pitoniak, 2006) was used for Sections I and II (multiple-choice questions). In this approach, for each question, a panelist decides on the likelihood (probability or chance) that a JQC would answer it correctly. Panelists made their judgments using the following rating scale: 0, .05, .10, .20, .30, .40, .50, .60, .70, .80, .90, .95, 1. The lower the value, the less likely it is that a JQC would answer the question correctly, because the question is difficult for the JQC. The higher the value, the more likely it is that a JQC would answer the question correctly.

For each panel, the panelists were asked to approach the judgment process in two stages. First, they reviewed the definition of the JQC and the question and decided if, overall, the question was difficult for the JQC, easy for the JQC, or moderately difficult/easy. The facilitator encouraged the panelists to consider the following rule of thumb to guide their decision:

- difficult questions for a JQC were in the 0 to .30 range;
- easy questions for a JQC were in the .70 to 1 range; and
- moderately difficult/easy questions for a JQC were in the .40 to .60 range.

The second decision was for panelists to decide how they wanted to refine their judgment within the range. For example, if a panelist thought that a question was easy for a JQC, the initial decision located the question in the .70 to 1 range. The second decision was for the panelist to decide if the likelihood of answering it correctly was .70, .80, .90, .95, or 1.0. The two-stage decision-process was implemented to reduce the cognitive load placed on the panelists. The panelists practiced making their standard-setting judgments on the first Listening set (six questions) in Section I.

The panelists engaged in two rounds of judgments. The Round 1 feedback provided to the panel included each panelist's (listed by ID number) recommended cut scores for Sections I and II (as well as cut scores for Sections III and IV) and the panel's average recommended cut score, highest and lowest cut score, and standard

deviation. Following discussion, the panelists' judgments were displayed for each question. The panelists' judgments were summarized by the three general difficulty levels (0 to .30, .40 to .60, and .70 to 1), and the panel's average question judgment was provided. Questions were highlighted to show when panelists converged in their judgments (approximately two-thirds of the panelists located a question in the same difficulty range) or diverged in their judgments. Panelists were asked to share their rationales for the judgments they made. Following this discussion, panelists were provided an opportunity to change their question-level standard-setting judgments (Round 2).

Other than the definition of the JQC, results from Panel 1 were not shared with the second panel. The question-level judgments and resulting discussions for Panel 2 were independent of judgments and discussions that occurred with Panel 1.

Standard Setting for Sections III and IV (Constructed-Response Questions). An Extended Angoff method (Cizek & Bunch, 2007; Hambleton & Plake, 1995) was used for Sections III and IV (constructed-response questions). In this approach, for each question, a panelist decides on the assigned score value that would most likely be earned by a JQC. The basic process that each panelist followed was first to review the definition of the JQC and then to review the question and the rubric for that question. The rubric for a question defines holistically the quality of the evidence that would merit a response earning a 3 (*High*), 2 (*Mid-High*), 1 (*Mid-Low*), or 0 (*Low*). During this review, each panelist independently considered the level of knowledge and/or skill required to respond to the question and the features of a response that would earn 3, 2, 1, or 0 points, as defined by the rubric.

A test taker's response to a constructed-response question is independently scored by two raters, and the sum of the raters' scores is the assigned score³; possible scores, therefore, range from zero (both raters assigned a score of zero) to six (both raters assigned a score of three). Each panelist decided on the score most likely to be earned by a JQC from the following possible values: 0, 1, 2, 3, 4, 5, 6. For each of the six constructed-response questions, panelists recorded the score (0 through 6) that a JQC would most likely earn. The panelists practiced making their standard-setting judgments on the first Writing question in Section III.

Consistent with the standard-setting process used for Sections I and II, the panelists engaged in two rounds of judgments for Sections III and IV. After the first round, the judgments of each panelist were summarized and projected for the panel to see and discuss. Each panelist's recommended cut score for Sections III and IV (as well as cut scores for Sections I and II) was displayed as was the panel's average recommended cut score, highest and lowest cut score, and standard deviation. The panelists' judgments also were displayed for each question. The

³ If the two raters' scores differ by more than one point (non-adjacent), the Chief Reader for that question assigns the score, which is then doubled.

panelists participated in a general discussion of the results. Panelists were asked to share their rationales for the judgments they made. Following this discussion, panelists were provided an opportunity to change their question-level standard-setting judgments (Round 2).

As with Sections I and II, results from Panel 1 were not shared with the second panel. The question-level judgments and resulting discussions for Panel 2 were independent of judgments and discussions that occurred with Panel 1.

Judgment of Praxis World Languages Content Specifications

Following the two-round standard setting process, each panel judged the importance of the knowledge and/or skills stated or implied in the assessment content specifications for the job of an entry-level K-12 teacher. The same content specifications were used to develop the German, French, and Spanish assessments. These judgments addressed the perceived content-based validity of the assessment. Judgments were made using a four-point Likert scale — *Very Important*, *Important*, *Slightly Important*, and *Not Important*. Each panelist independently judged the 21 knowledge/skills statements. (See Appendix G for the common content specifications for the German, French, and Spanish assessments.)

Results

Initial Evaluation Forms

The panelists completed two initial evaluation forms, once after they were trained in how to make their standard-setting judgments for Sections I and II (multiple-choice questions), and once after they were trained to make their judgments for Sections III and IV (constructed-response questions). The primary information collected from these forms was the panelists indicating if they had received adequate training to make their standard-setting judgments and were ready to proceed. Across all assessments and panels, all panelists indicated that they were prepared to make their judgments.

Summary of Standard Setting Judgments by Round

A summary of each round of standard-setting judgments for Sections I and II (multiple-choice questions), Sections III and IV (constructed-response questions), and the overall assessment is presented in Appendix D (German), Appendix E (French), and Appendix F (Spanish). The numbers in each table reflect the recommended cut scores — the number of raw points needed to “pass” the section or test — of each panelist for the two rounds. Note that the Praxis World Languages assessments report a single, overall score and that the panels are recommending a single cut score for the combination of Sections I, II, III and IV. The separate “cut scores” for the four sections are intermediate steps in calculating the overall cut score. For each assessment, the panels’ average recommended cut score and highest and lowest cut scores are reported, as are the standard deviations (SD) of

panelists' cut scores and the standard errors of judgment (SEJ). The SEJ is one way of estimating the reliability of the judgments. It indicates how likely it would be for other panels of educators similar in make-up, experience, and standard-setting training to the current panels to recommend the same cut score on the same form of the test. A comparable panel's cut score would be within 1 SEJ of the current average cut score 68 percent of the time and within 2 SEJs 95 percent of the time.

Round 1 judgments are made without discussion among the panelists. The most variability in judgments, therefore, is typically present in the first round. Round 2 judgments, however, are informed by panel discussion; thus, it is common to see a decrease both in the standard deviation and SEJ. This decrease — indicating convergence among the panelists' judgments — was observed for four of the six panels; the standard deviation increased somewhat between rounds for the first German and Spanish panels.

For each assessment, the Round 2 average score for each section is summed to arrive at each panel's overall recommended cut score (passing score). It should be noted, however, that there are no required minimum section scores that must be obtained in order to pass the German, French, or Spanish assessments. The total test cut score is compensatory, in that as long as the total cut score is met or exceeded, the candidate has passed

Praxis German Assessment

The panels' cut score recommendations for the Praxis German assessment are 65.71 for Panel 1 and 62.09 for Panel 2 (see Tables 2a and 3a in Appendix D). The values were rounded to the next highest whole number to determine the functional recommended cut scores — 66 for Panel 1 and 63 for Panel 2. The values of 66 and 63 represent approximately 67% and 64%, respectively, of the total available 98 raw points that could be earned on the assessment. The scaled scores associated with 66 and 63 raw points are 165 and 161, respectively.⁴

Tables 4a and 4b (in Appendix D) present the estimated standard errors of measurement (SEM) around the recommended cut scores for each panel. A standard error represents the uncertainty associated with a test score. The scaled scores associated with 1 and 2 SEMs are provided. The standard errors provided are an estimate, given that the Praxis German assessment has not yet been administered.

The recommended cut scores for each panel, as well as the average cut across the two panels, are provided to help state departments of education determine an appropriate cut (or passing) score for the Praxis German assessment. The panels' average cut score recommendation for the Praxis German assessment is 63.90. The value was rounded to 64 (next highest raw score) to determine the functional recommended cut score. The value of 64 represents approximately 65% of the total available 98 raw points that could be earned on the assessment.

⁴ For reference purposes, if the recommended raw cut score were 65 or 62 points, the scaled score would be 164 or 160, respectively.

The scaled score associated with 64 raw points is 163.⁵ Table 4c (in Appendix D) presents the standard error of measurement (SEM) around the recommended cut score combining the information from the two panels.

Praxis French Assessment

The panels' cut score recommendations for the Praxis French assessment are 58.54 for Panel 1 and 65.84 for Panel 2 (see Tables 8a and 9a in Appendix E). The values were rounded to the next highest whole number to determine the functional recommended cut scores — 59 for Panel 1 and 66 for Panel 2. The values of 59 and 66 represent approximately 61% and 68%, respectively, of the total available 97 raw points that could be earned on the assessment. The scaled scores associated with 59 and 66 raw points are 157 and 166, respectively.⁶

Tables 10a and 10b (in Appendix E) present the estimated standard errors of measurement (SEM) around the recommended cut scores for each panel. A standard error represents the uncertainty associated with a test score. The scaled scores associated with 1 and 2 SEMs are provided. The standard errors provided are an estimate, given that the Praxis French assessment has not yet been administered.

The recommended cut scores for each panel, as well as the average cut across the two panels, are provided to help state departments of education determine an appropriate cut (or passing) score for the Praxis French assessment. The panels' average cut score recommendation for the Praxis French assessment is 62.19. The value was rounded to 63 (next highest raw score) to determine the functional recommended cut score. The value of 63 represents approximately 65% of the total available 97 raw points that could be earned on the assessment. The scaled score associated with 63 raw points is 162.⁷ Table 10c (in Appendix E) presents the standard error of measurement (SEM) around the recommended cut score combining the information from the two panels.

Praxis Spanish Assessment

The panels' cut score recommendations for the Praxis Spanish assessment are 65.54 for Panel 1 and 68.02 for Panel 2 (see Tables 14a and 15a in Appendix F). The values were rounded to the next highest whole number to determine the functional recommended cut scores — 66 for Panel 1 and 69 for Panel 2. The values of 66 and 69 represent approximately 69% and 72%, respectively, of the total available 96 raw points that could be earned on the assessment. The scaled scores associated with 66 and 69 raw points are 167 and 171, respectively.⁸

Tables 16a and 16b (in Appendix F) present the estimated standard errors of measurement (SEM) around the recommended cut scores for each panel. A standard error represents the uncertainty associated with a test score.

⁵ For reference purposes, if the recommended raw cut score was 63 points, the scaled score would be 161.

⁶ For reference purposes, if the recommended raw cut score were 58 or 65 points, the scaled score would be 156 or 165, respectively.

⁷ For reference purposes, if the recommended raw cut score was 62 points, the scaled score would be 161.

⁸ For reference purposes, if the recommended raw cut score were 65 or 68 points, the scaled score would be 166 or 170, respectively.

The scaled scores associated with 1 and 2 SEMs are provided. The standard errors provided are an estimate, given that the Praxis Spanish assessment has not yet been administered.

The recommended cut scores for each panel, as well as the average cut across the two panels, are provided to help state departments of education determine an appropriate cut (or passing) score for the Praxis Spanish assessment. The panels' average cut score recommendation for the Praxis Spanish assessment is 66.78. The value was rounded to 67 (next highest raw score) to determine the functional recommended cut score. The value of 67 represents approximately 70% of the total available 96 raw points that could be earned on the assessment. The scaled score associated with 67 raw points is 168.⁹ Table 16c (in Appendix F) presents the standard error of measurement (SEM) around the recommended cut score combining the information from the two panels.

Summary of Specification Judgments

Panelists judged the extent to which the knowledge and/or skills reflected by the Praxis World Languages assessments content specifications were important for entry-level teachers. Panelists rated the 21 knowledge/skills statements on a four-point scale ranging from *Very Important* to *Not Important*. The panelists' ratings are summarized in Table 5 (in Appendix D) for German, Table 11 (in Appendix E) for French, and Table 17 (in Appendix F) for Spanish.

Across the three assessment, only one knowledge/skills statement — “*Knows how to move beyond literal comprehension in the interpretive mode (listening) by inferring ...*” — was judged to be *Very Important* or *Important* by less than 75% of the panelists for a particular language, German. Two knowledge/skills statements were judged to be *Very Important* or *Important* by less than 90% of the panelists for two languages:

- “*Knows how to communicate orally in the presentational mode (speaking) by delivering oral presentations on familiar literary or cultural topics ...*” for German and Spanish; and
- “*Knows how to contrast syntactical patterns of simple sentences and questions with those of English*” for French and Spanish.

The complete texts of the content specifications are presented in Appendix G.

⁹ For reference purposes, if the recommended raw cut score was 66 points, the scaled score would be 167.

Summary

To support the decision-making process for state departments of education with regards to establishing passing scores, or cut scores, for the Praxis World Languages assessments for German, French, and Spanish, research staff from Educational Testing Service (ETS) designed and conducted a series of multi-state standard setting studies. The studies also collected content-related validity evidence to confirm the importance of the content specifications for entry-level K-12 German, French, and Spanish teachers. The standard setting studies involved two expert panels for each assessment, comprised of teachers, administrators, and college faculty.

Standard setting was conducted using a probability-based Angoff approach (for the multiple-choice sections) and an Extended Angoff approach (for the constructed-response sections). Section-level minimum scores were constructed and an overall cut score was computed. The recommended cut scores for each panel, as well as the average cut across the two panels, are provided to help state departments of education determine appropriate cut (or passing) scores.

- For Praxis World Languages: **German**, the average recommended cut score is **64** (on the raw score metric), which represents 65% of total available 98 raw points (the recommended cut scores for Panels 1 and 2 are 66 and 63, respectively). The scaled score associated with a raw score of 64 on the Praxis German assessment is 163.
- For Praxis World Languages: **French**, the average recommended cut score is **63** (on the raw score metric), which represents 65% of total available 97 raw points (the recommended cut scores for Panels 1 and 2 are 59 and 66, respectively). The scaled score associated with a raw score of 63 on the Praxis French assessment is 162.
- For Praxis World Languages: **Spanish**, the recommended cut score is **67** (on the raw score metric), which represents 70% of total available 96 raw points (the recommended cut scores for Panels 1 and 2 are 66 and 69, respectively). The scaled score associated with a raw score of 67 on the Praxis Spanish assessment is 168.

For each of the assessments, both panels confirmed that the knowledge and/or skills stated or implied in the Praxis World Languages assessment content specifications were important for entry-level K-12 teachers. The results of the evaluation surveys (initial and final) from each panels support the quality of the standard-setting implementation.

References

- Brandon, P.R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education*, 17, 59-88.
- Cizek, G. J., & Bunch, M.B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- Educational Testing Service. (in press). *German: World Languages: Test at a glance*. Princeton, NJ: Author.
- Educational Testing Service. (in press). *French: World Languages: Test at a glance*. Princeton, NJ: Author.
- Educational Testing Service. (in press). *Spanish: World Languages: Test at a glance*. Princeton, NJ: Author.
- Hambleton, R. K., & Pitoniak, M.J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational Measurement* (4 ed., pp. 433-470). Westport, CT: American Council on Education/Praeger.
- Hambleton, R. K., & Plake, B.S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education*, 8, 41-55.

APPENDIX A
Common Workshop Agenda

AGENDA

Praxis [Target Language]: World Languages Assessment

Standard Setting Study

Day 1

8:00 – 8:15	Welcome and Introduction
8:15 – 8:45	Overview of Standard Setting & Workshop Events
8:45 – 9:15	Overview of the Praxis World Languages Assessments
9:15 – 9:20	Break
9:20 – 11:30	“Take” the Praxis [Target Language]: World Languages Assessment
11:30 – 12:00	Discuss the Praxis [Target Language]: World Languages Assessment
12:00 – 12:15	Define the Knowledge/Skills of a JQC
12:15 – 1:00	Lunch
1:00 – 3:00	Define the Knowledge/Skills of a JQC (continued)
3:00 – 3:15	Break
3:15 – 3:45	Standard Setting Training for M-C Questions (Sections I and II)
3:45 – 5:15	Round 1 Standard Setting Judgments for Multiple-Choice
5:15 – 5:30	Collect Materials; End of Day 1

AGENDA

Praxis [Target Language]: World Languages Assessment

Standard Setting Study

Day 2

9:00 – 9:15	Questions from Day 1 & Overview of Day 2
9:15 – 10:00	Standard Setting Training for CR Questions (Sections III and IV)
10:00 – 10:30	Round 1 Standard Setting Judgments for Constructed-Response
10:30 – 10:35	Break
10:35 – 12:00	Round 1 Feedback & Round 2 Judgments
12:00 – 12:45	Lunch
12:45 – 2:15	Round 1 Feedback & Round 2 Judgments (continued)
2:15 – 3:00	Specification Judgment
3:00 – 3:15	Feedback on Round 2 Recommended Cut Score
3:15 – 3:30	Complete Final Evaluation
3:30 – 3:45	Collect Materials; End of Study

APPENDIX B
Just Qualified Candidate (JQC) Definitions

Definition of the Just Qualified Candidate – German

Listening, Reading, and Cultural Knowledge

1. Ability to use basic reading strategies, such as word analysis, inference, and context clues, with authentic samples/materials
2. Have a rich, passive German vocabulary which includes high-frequency idioms and grammatical terminology
3. Comprehend a reasonable amount of main ideas, key concepts and some details in authentic samples of paragraph-length discourse
4. In aural and written communication, recognizes various registers and voices to facilitate comprehension
5. Has a basic understanding of syntactical relationships and major verb tenses and moods
6. Can distinguish between phonemes and diphthongs
7. Generally identify significant current, historical, and/or cultural people, places, events, and social structures in German-speaking countries
8. Has a basic understanding of regional differences in language

Writing and Speaking

1. Ability to adjust pace, intonation, and fluency of delivery
2. Is able to be comprehensible to a native speaker through articulation and pronunciation
3. Can express himself/herself on a variety of concrete and factual topics
4. Has a diverse active vocabulary which allows them to successfully circumlocute, summarize and paraphrase
5. Demonstrates control of mechanics and conventions in writing
6. Demonstrates control of conventions in discourse
7. Is able to adjust writing and speaking for various purposes and audiences
8. Is able to sequence ideas and use conjunctions and transitions to achieve cohesion in writing

Definition of the Just Qualified Candidate – French

Listening, Reading, and Cultural Knowledge

1. Uses basic reading strategies such as word analysis, inference, and context clues with authentic texts
2. Comprehends a broad French vocabulary including idioms
3. Comprehends (a) main ideas, (b) most key concepts and (c) some details in authentic aural and written communication
4. Recognizes various registers and formal/informal voices to facilitate comprehension in authentic aural and written communication
5. Has an understanding of grammar, including syntax, major verb tenses and moods
6. Has a basic knowledge of French pronunciation
7. Can identify historical or current people, places, events, and social structures in French-speaking countries or regions
8. Has a basic awareness of regional differences in language

Writing and Speaking

1. Is comprehensible to a native speaker.
2. Can express himself/herself on a variety of concrete and factual topics, including personal opinions
3. Uses a diverse vocabulary to circumlocute, summarize and paraphrase successfully in writing and speaking and engage in conversations
4. Demonstrates basic command of mechanics and conventions in writing
5. Demonstrates control of conventions in speaking
6. Adjusts writing and speaking for various purposes and audiences
7. Sequences ideas to achieve cohesion in writing and speaking

Definition of the Just Qualified Candidate – Spanish

Listening, Reading, and Cultural Knowledge

1. Uses basic reading strategies such as word analysis, inference, and context clues with authentic texts
2. Comprehends a broad Spanish vocabulary including widely used idiomatic expressions
3. Comprehends (a) main ideas, (b) most subordinate ideas and (c) some details in authentic aural and written communication
4. Comprehends meanings of various registers and formal/informal voice in authentic aural and written communication
5. Has an understanding of grammar, including syntax, verb tenses and moods
6. Has a general knowledge of Spanish pronunciation
7. Can identify historical or current people, places, events, and social structures in Spanish-speaking countries or regions
8. Has a basic awareness of regional differences in language

Writing and Speaking

1. Is comprehensible to a native speaker.
2. Can express himself/herself on a variety of concrete and factual topics, and express and defend personal opinions
3. Uses a diverse vocabulary to circumlocute, summarize and paraphrase successfully in writing and speaking
4. Appropriately applies mechanics and conventions in writing and speaking
5. Writes and speaks appropriately for various purposes and to varied audiences
6. Sequences ideas to achieve cohesion in writing and speaking

APPENDIX C
Panelists' Names & Affiliations

German Panel 1

Panelist

Sandra Achenbach
 Amy L. Bauer
 James H. Bright
 Mary Ann Crow
 Stephanie Draheim
 Christi Elkins-Gabbard
 VidaJane Haynes
 Brad Martin
 Erin McKeag
 Susan Peterson
 Colleen Richards
 Claudia Schoellkopf
 Wiebke Strehl
 Shauna Winegar
 Maga Isabel Wisard

Affiliation

Hardin Valley Academy, Knox County School (TN)
 Rapid City Central High School (SD)
 Henry Clay High School, Fayette County Public Schools (KY)
 Bismarck High School (ND)
 Menasha Joint School District (WI)
 Fayette County Schools (KY)
 McGavock Comprehensive High School (TN)
 Elkins High School (WV)
 Charlotte Mecklenburg Schools (NC)
 T.F. Riggs High School, South Dakota District 32-2 (SD)
 Butler Area School District (PA)
 Bismarck Public Schools (ND)
 University of South Carolina (SC)
 Mt. Crest High School, Cache County School District (UT)
 Poplarville Elementary School (MS)

German Panel 2

Panelist

Anthony M.DeRosa
 Donna M. Evans
 J. Sarah Floyd
 Sarah Glasser
 Melissa Hadorn
 Arthur D. Holder
 Diana T. Ihlenfeld
 Susanne Lenné Jones
 Elke K. Kuegle
 Joy E. Loomis
 Joan S. MacDonald
 Michelle Mattson
 Cody Mickelson
 Michael C. Netzloff
 Andrew J. Richards
 Dorothée Rosser
 Annette Sherrer

Affiliation

Thomas S. Wootton High (MD)
 Las Vegas Academy/Clark County School District (NV)
 Lexington High School (SC)
 Wright Jr/Sr High (WY)
 Sturgis Brown High School (SD)
 Judge Memorial Catholic High School (UT)
 Ohio County Schools (WV)
 East Carolina University (NC)
 Stevens High School, Rapid City Area Schools (SD)
 Newark High School (DE)
 Martin Luther King Magnet (TN)
 Rhodes College (TN)
 Jamestown Public School District #1 (ND)
 Bismarck Public Schools (ND)
 Fox Chapel Area School District (PA)
 Gadsden City High School (AL)
 Picayune Memorial High School (MS)

French Panel 1

Panelist

Anita J. Alkhas
 Pierre C. Baigue
 Colette Ballew
 Claudia V. Bezaka
 Paula Summers Calderon
 Cristina Carlotti
 Stephen M. Dubrow
 Nancy Erickson
 Gail Fahy
 Antoine F.Gnintedem
 Melissa Hadorn
 Sherri K. Harkins
 Leanne Hinkle
 Wendy D. Howard
 Elisabeth Kohl
 William Mann
 Shawn Morrison
 Oscar Niyiragira
 Anne Olafson
 Amanda Robustelli-Price
 Jacquelyn Sergi
 William Thompson
 Jocelyn A. M. Waddle

Affiliation

University of Wisconsin-Milwaukee (WI)
 Granite School District (UT)
 Wayne Highlands School District (PA)
 District of Columbia Public Schools (DC)
 Louisiana State University and A&M College (LA)
 East Providence High School (RI)
 Walter Johnson High School (MD)
 University of Southern Maine (ME)
 Palo Verde HS Clark County School District (NV)
 Sunflower County School District (MS)
 Sturgis Brown High School (SD)
 Wicomico County Public Schools (MD)
 Bolton High School (TN)
 Gaston County School District (NC)
 Council Rock High School –South (PA)
 Clay County High School (WV)
 College of Charleston (SC)
 Jefferson County Public Schools (KY)
 Minot High School (ND)
 Bristol Central High School (CT)
 South Panola High School (MS)
 The University of Memphis (TN)
 Frankfort High School (KY)

French Panel 2

Panelist

Lydia Wilson Kohler
 Robert Desmarais Sullivan
 Denise B. Benskin
 Crecia C. Swaim
 Jason Bagley
 Mary C. Frye
 Mary Anne Smith
 Robert Denis
 Nancy Jarchow
 Madeleine Hooper-Kernen
 Nancy P. Wilson
 Robert G. Erickson
 Elizabeth Howe
 Suzanne Lord Guazzoni
 Timothy Wung Kum
 Stephanie Viator
 Wendy C. Mumy
 Jan Hennessey
 Tracy Lambert
 Stephen Keller
 Margaret Schmidt Dess
 J. Karine Simpson
 Linda E. Lassiter
 Valerie Kling

Affiliation

George Rogers Clark High School (KY)
 Hattiesburg High School (MS)
 Prince Georges County (MD)
 Betsey Ross Arts Magnet School (CT)
 Lexington High School (SC)
 West Virginia State University (WV)
 Pearl City High School (HI)
 Las Vegas High School (NV)
 Williamstown High School (VT)
 Missouri State University (MO)
 Mifflin School District (PA)
 Brigham Young University (UT)
 Hardin Valley Academy (TN)
 Stone High School (VT)
 Greenville-Weston High School (MS)
 Cedar Creek School (LA)
 West Craven High School (NC)
 Dover High School (NH)
 Lafayette High School (KY)
 A.C. Flora High School (SC)
 Shorewood High School (WI)
 Central Bucks School District (PA)
 Southern University and A&M College (LA)
 Bismarck High School (ND)

Spanish Panel 1

Panelist

Ignacio M. Cariaga
 June C. D. Carter
 Eric O. Cintrón
 Larissa Cuevas
 Stephanie Dominguez
 Paul Fallon
 Geoffrey Gillett
 Bridget Suárez Kalmar
 José Labrado
 Mina T. Levenson
 Terri Marlow
 Belgica Nina-Matos
 Samuel J. Ogdie
 Lisa Ramey
 Joyce Richburg
 Ruth E. Smith
 Nancy E. Yetter
 Thomasina I. White

Affiliation

State of Hawaii Public Schools (HI)
 University of South Carolina Upstate (SC)
 Plymouth State University (NH)
 Pass Christian School District (MS)
 Smithville R-II School District (MO)
 East Carolina University (NC)
 Maine School Administrative District 41 (ME)
 Craftsbury Schools (VT)
 Dawson Springs High School (KY)
 Pittsburgh Public Schools (PA)
 Wood County Schools (WV)
 Delmar School District (DE)
 Augustana College (SD)
 North Central Public School (ND)
 Birmingham City Schools (AL)
 University of Louisiana Monroe (LA)
 Baltimore County Public Schools (MD)
 School District of Philadelphia (PA)

Spanish Panel 2

Panelist

Carolyn A. Anderson
 Isabel Cavour
 Angela Culver Johnson
 Telece Marbrey
 Luis M. González-García
 Sharon M. Gracia
 Marta C. Gumpert
 Andrés V. Hernández
 David Herren
 Grace Leavitt
 Jennifer Love
 Raquel Oxford
 Nancy S. Ryan
 Ángel T. Tuninetti
 Diane VanDenOever
 Summer Van Wagnen
 Isabel Vázquez-Gil
 Nancy Wahineokai
 Giovanna Yaranga-Reyes
 James R. Yoder
 Dina Zavala-Petherbridge

Affiliation

Barnwell School District #45 (SC)
 University of Dayton (OH)
 Madison City Schools (AL)
 Knox County Schools (TN)
 Northern Kentucky University (KY)
 Granite School District (UT)
 Southeastern Louisiana University (LA)
 Biloxi Public Schools (MS)
 Union High School (VT)
 Greely High School / St. Joseph's College (ME)
 Prince George's County Public Schools (MD)
 University of Wisconsin Milwaukee (WI)
 Berkeley County West Virginia Schools (WV)
 West Virginia University (WV)
 The University of Sioux Falls (SD)
 Wake County Public School System (NC)
 District of Columbia Public Schools (DC)
 Radford High School (HI)
 Burlington School District (VT)
 Clark County School District (NV)
 Valley City State University (ND)

APPENDIX D
Results for Praxis World Languages: German

Table 1 Committee Member Demographics — German

	Panel 1		Panel 2	
	N	Percent	N	Percent
Group you are representing				
Teachers	14	93%	14	82%
Administrator/Department Head	0	0%	1	6%
College Faculty	1	7%	2	12%
Other	0	0%	0	0%
Race				
African American or Black	0	0%	0	0%
Alaskan Native or American Indian	0	0%	0	0%
Asian or Asian American	0	0%	0	0%
Native Hawaiian or Other Pacific Islander	0	0%	0	0%
White	14	93%	17	100%
Hispanic	1	7%	0	0%
Gender				
Female	13	87%	12	71%
Male	2	13%	5	29%
In which language are you most fluent?				
English	9	60%	12	71%
German	0	0%	0	0%
English and German about the same	5	33%	5	29%
Other	1	7%	0	0%
Are you certified as a German teacher in your state?				
No	1	7%	2	12%
Yes	14	93%	15	88%
Are you currently teaching German in your state?				
No	1	7%	1	6%
Yes	14	93%	16	94%
Are you currently mentoring another German teacher?				
No	14	93%	16	94%
Yes	1	7%	1	6%
How many years of experience do you have as a German teacher in your state?				
3 years or less	1	7%	3	18%
4 - 7 years	4	27%	4	24%
8 - 11 years	4	27%	2	12%
12 - 15 years	2	13%	3	18%
16 years or more	4	27%	5	29%
For which education level are you currently teaching German?				
Elementary (K - 5 or K - 6)	1	7%	0	0%
Middle School (6 - 8 or 7 - 9)	2	13%	0	0%
High School (9 - 12 or 10 - 12)	9	60%	14	82%
Middle/High School	1	7%	1	6%
Higher Education	1	7%	2	12%
Other	1	7%	0	0%
School Setting				
Urban	8	53%	6	35%
Suburban	3	20%	7	41%
Rural	4	27%	4	24%

Table 2a Cut score Summary by Round of Judgments — German Panel 1

	Round 1	Round 2
Section I: Listening (Max. Raw Score = 30)		
Average	20.74	20.51
SD	1.97	2.05
SEJ	0.51	0.53
Highest	23.70	23.20
Lowest	15.45	14.80
Section II: Reading (Max. Raw Score = 32)		
Average	23.31	22.67
SD	1.43	1.41
SEJ	0.37	0.36
Highest	26.05	25.00
Lowest	20.85	19.10
Section III: Writing (Max. Raw Score = 18)		
Average	11.67	11.33
SD	1.45	1.05
SEJ	0.37	0.27
Highest	15.00	13.00
Lowest	10.00	10.00
Section IV: Speaking (Max. Raw Score = 18)		
Average	11.60	11.20
SD	2.29	2.01
SEJ	0.59	0.52
Highest	15.00	14.00
Lowest	7.00	7.00
Total (Max. Raw Score = 98)		
Average	67.32	65.71
SD	5.17	5.84
SEJ	1.34	1.51
Highest	76.90	74.20
Lowest	56.75	50.90

Table 2b Panelists Cut scores by Round of Judgments — German Panel 1

Panelist	Section I		Section II		Section III		Section IV		Total	
	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2
1	20.40	20.60	23.80	23.30	14.00	12.00	13.00	12.00	71.20	67.90
2	20.40	20.05	22.40	22.00	11.00	11.00	12.00	12.00	65.80	65.05
3	20.75	21.05	23.30	23.40	11.00	11.00	11.00	12.00	66.05	67.45
4	21.40	21.90	22.95	23.85	13.00	13.00	13.00	13.00	70.35	71.75
5	23.70	23.20	25.35	25.00	13.00	13.00	14.00	13.00	76.05	74.20
6	15.45	14.80	21.30	19.10	11.00	10.00	9.00	7.00	56.75	50.90
7	19.50	18.50	24.80	22.80	12.00	10.00	7.00	8.00	63.30	59.30
8	19.80	19.90	23.00	22.90	10.00	11.00	9.00	9.00	61.80	62.80
9	21.85	20.90	23.60	22.20	11.00	11.00	12.00	12.00	68.45	66.10
10	21.10	20.50	22.45	22.20	11.00	11.00	12.00	12.00	66.55	65.70
11	23.00	23.20	23.90	24.10	15.00	13.00	15.00	13.00	76.90	73.30
12	21.60	19.40	24.00	21.85	10.00	10.00	10.00	10.00	65.60	61.25
13	18.75	20.45	21.95	22.65	11.00	12.00	15.00	14.00	66.70	69.10
14	20.85	20.70	26.05	23.65	11.00	11.00	12.00	11.00	69.90	66.35
15	22.60	22.45	20.85	21.05	11.00	11.00	10.00	10.00	64.45	64.50

Table 3a Cut score Summary by Round of Judgments — German Panel 2

	Round 1	Round 2
Section I: Listening (Max. Raw Score = 30)		
Average	18.48	18.09
SD	2.36	2.00
SEJ	0.57	0.48
Highest	23.55	22.65
Lowest	13.60	14.20
Section II: Reading (Max. Raw Score = 32)		
Average	21.16	21.00
SD	2.28	1.86
SEJ	0.55	0.45
Highest	26.75	24.60
Lowest	17.45	17.35
Section III: Writing (Max. Raw Score = 18)		
Average	12.35	12.12
SD	1.11	1.22
SEJ	0.27	0.30
Highest	14.00	14.00
Lowest	10.00	10.00
Section IV: Speaking (Max. Raw Score = 18)		
Average	11.35	10.88
SD	1.62	1.11
SEJ	0.39	0.27
Highest	15.00	13.00
Lowest	8.00	9.00
Total (Max. Raw Score = 98)		
Average	63.34	62.09
SD	4.47	4.11
SEJ	1.08	1.00
Highest	73.30	69.25
Lowest	56.70	53.05

Table 3b Panelists Cut scores by Round of Judgments — German Panel 2

Panelist	Section I		Section II		Section III		Section IV		Total	
	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2
1	18.50	17.70	23.80	23.10	13.00	13.00	13.00	13.00	68.30	66.80
2	22.20	20.10	23.40	22.20	13.00	13.00	10.00	10.00	68.60	65.30
3	20.25	20.05	19.90	21.10	13.00	13.00	11.00	11.00	64.15	65.15
4	18.65	18.25	18.70	18.75	11.00	10.00	11.00	11.00	59.35	58.00
5	19.90	20.20	24.00	23.70	10.00	10.00	10.00	10.00	63.90	63.90
6	18.30	18.10	20.10	19.20	14.00	12.00	13.00	12.00	65.40	61.30
7	18.95	18.05	20.75	20.45	12.00	12.00	11.00	10.00	62.70	60.50
8	16.90	15.70	17.45	17.35	12.00	11.00	11.00	9.00	57.35	53.05
9	23.55	22.65	26.75	24.60	11.00	11.00	12.00	11.00	73.30	69.25
10	18.40	18.20	21.40	21.10	13.00	14.00	12.00	10.00	64.80	63.30
11	18.20	18.00	21.15	22.65	14.00	14.00	8.00	10.00	61.35	64.65
12	17.75	17.75	21.55	21.25	13.00	13.00	12.00	11.00	64.30	63.00
13	19.20	19.20	19.80	20.50	13.00	13.00	15.00	13.00	67.00	65.70
14	13.60	14.20	19.00	19.30	13.00	12.00	12.00	11.00	57.60	56.50
15	18.55	16.45	21.50	20.60	11.00	11.00	11.00	11.00	62.05	59.05
16	16.10	16.95	19.90	19.95	12.00	12.00	12.00	12.00	60.00	60.90
17	15.20	16.00	20.50	21.20	12.00	12.00	9.00	10.00	56.70	59.20

Table 4a Cut scores within 1 and 2 SEMs of the Recommended Cut score — German Panel 1

Recommended Cut score (SEM)		Scale Score Equivalent
66 (4.50)		165
- 2 SEMs	57	153
-1 SEM	62	160
+1 SEM	71	172
+ 2 SEMs	75	177

Table 4b Cut scores within 1 and 2 SEMs of the Recommended Cut score — German Panel 2

Recommended Cut score (SEM)		Scale Score Equivalent
63 (4.66)		161
- 2 SEMs	53	148
-1 SEM	58	155
+1 SEM	67	166
+ 2 SEMs	72	173

Table 4c Cut scores within 1 and 2 SEMs of the Recommended Cut score — Combined German Panels

Recommended Cut score (SEM)		Scale Score Equivalent
64 (4.59)		163
- 2 SEMs	55	151
-1 SEM	60	157
+1 SEM	69	169
+ 2 SEMs	74	175

Note: Consistent with the recommended cut score, the cut scores at the different SEMs have been rounded to the next highest whole number.

Table 5 Specification Judgments — German (Panels 1 & 2 Judgments Combined)

	Very Important		Important		Slightly Important		Not Important	
	N	%	N	%	N	%	N	%
Language, Linguistics, and Comparison								
A. Demonstrating Language Proficiency	27	84%	5	16%	0	0%	0	0%
Subtopic 1	25	78%	7	22%	0	0%	0	0%
Subtopic 2	25	78%	7	22%	0	0%	0	0%
Subtopic 3	14	44%	18	56%	0	0%	0	0%
Subtopic 4	15	47%	16	50%	1	3%	0	0%
Subtopic 5	19	59%	13	41%	0	0%	0	0%
Subtopic 6	21	66%	9	28%	2	6%	0	0%
Subtopic 7	5	16%	19	59%	8	25%	0	0%
Subtopic 8	7	22%	23	72%	2	6%	0	0%
Subtopic 9	27	84%	5	16%	0	0%	0	0%
Subtopic 10	10	31%	22	69%	0	0%	0	0%
Subtopic 11	15	47%	12	38%	4	13%	1	3%
B. Understanding Linguistics	18	60%	11	37%	1	3%	0	0%
Subtopic 1	20	63%	12	38%	0	0%	0	0%
Subtopic 2	17	53%	13	41%	2	6%	0	0%
Subtopic 3	17	53%	14	44%	1	3%	0	0%
Subtopic 4	20	63%	12	38%	0	0%	0	0%
Subtopic 5	27	84%	4	13%	1	3%	0	0%
Subtopic 6	19	59%	12	38%	1	3%	0	0%
C. Comparison of Target Language with English	13	42%	14	45%	4	13%	0	0%
Subtopic 1	13	41%	16	50%	3	9%	0	0%
Subtopic 2	14	44%	16	50%	2	6%	0	0%
Cultures, Literature, Cross-disciplinary Concepts								
A. Demonstrating Language Proficiency	16	52%	15	48%	0	0%	0	0%
Subtopic 1	20	65%	10	32%	1	3%	0	0%
Subtopic 2	12	39%	15	48%	4	13%	0	0%

Table 6a Final Evaluation — German Panel 1

	Strongly Agree		Agree		Disagree		Strongly Disagree	
	N	%	N	%	N	%	N	%
I understood the purpose of this study.	15	100%	0	0%	0	0%	0	0%
The instructions and explanations provided by the facilitator were clear.	14	93%	1	7%	0	0%	0	0%
The training in the standard setting methods was adequate to give me the information I needed to complete my assignment.	15	100%	0	0%	0	0%	0	0%
The explanation of how the recommended cut scores are computed was clear.	13	87%	2	13%	0	0%	0	0%
The opportunity for feedback and discussion between rounds was helpful.	15	100%	0	0%	0	0%	0	0%
The process of making the standard setting judgments was easy to follow.	9	60%	6	40%	0	0%	0	0%
	Very Influential		Somewhat Influential		Not Influential			
How influential was each of the following factors in guiding your standard setting judgments?	N	%	N	%	N	%		
The definition of the Just Qualified Candidate	15	100%	0	0%	0	0%		
The between-round discussions	11	73%	4	27%	0	0%		
The cut scores of other panel members	4	27%	7	47%	4	27%		
My own professional experience	10	67%	5	33%	0	0%		
	Very Comfortable		Somewhat Comfortable		Somewhat Uncomfortable		Very Uncomfortable	
	N	%	N	%	N	%	N	%
Overall, how comfortable are you with the panel's recommended cut score? ¹⁰								
	Too Low		About Right		Too High			
Overall, the panel's recommended cut score is: ¹⁰								

¹⁰ Due to technical problems during the study, panelists were not able to review and judge their comfort level with the overall cut score following Round 2.

Table 6b Final Evaluation — German Panel 2

	Strongly Agree		Agree		Disagree		Strongly Disagree	
	N	%	N	%	N	%	N	%
I understood the purpose of this study.	16	94%	1	6%	0	0%	0	0%
The instructions and explanations provided by the facilitator were clear.	13	76%	4	24%	0	0%	0	0%
The training in the standard setting methods was adequate to give me the information I needed to complete my assignment.	14	82%	3	18%	0	0%	0	0%
The explanation of how the recommended cut scores are computed was clear.	15	88%	2	12%	0	0%	0	0%
The opportunity for feedback and discussion between rounds was helpful.	14	82%	3	18%	0	0%	0	0%
The process of making the standard setting judgments was easy to follow.	9	53%	7	41%	1	6%	0	0%
	Very Influential		Somewhat Influential		Not Influential			
How influential was each of the following factors in guiding your standard setting judgments?	N	%	N	%	N	%		
The definition of the Just Qualified Candidate	14	82%	3	18%	0	0%		
The between-round discussions	11	65%	2	12%	4	24%		
The cut scores of other panel members	2	12%	9	53%	6	35%		
My own professional experience	8	47%	8	47%	1	6%		
	Very Comfortable		Somewhat Comfortable		Somewhat Uncomfortable		Very Uncomfortable	
	N	%	N	%	N	%	N	%
Overall, how comfortable are you with the panel's recommended cut score?	10	59%	6	35%	1	6%	0	0%
	Too Low		About Right		Too High			
Overall, the panel's recommended cut score is:	1	6%	16	94%	0	0%		

APPENDIX E
Results for Praxis World Languages: French

Table 7 Committee Member Demographics — French

	Panel 1		Panel 2	
	N	Percent	N	Percent
Group you are representing				
Teachers	15	65%	19	79%
Administrator/Department Head	2	9%	2	8%
College Faculty	5	22%	2	8%
Other	1	4%	1	4%
Race				
African American or Black	3	13%	3	13%
Alaskan Native or American Indian	1	4%	0	0%
Asian or Asian American	0	0%	1	4%
Native Hawaiian or Other Pacific Islander	0	0%	0	0%
White	19	83%	19	79%
Hispanic	0	0%	0	0%
Gender				
Female	17	74%	18	75%
Male	6	26%	6	25%
In which language are you most fluent?				
English	14	61%	19	79%
French	1	4%	3	13%
English and French about the same	7	30%	2	8%
Other	1	4%	0	0%
Are you certified as a French teacher in your state?				
No	4	17%	4	17%
Yes	19	83%	20	83%
Are you currently teaching French in your state?				
No	2	9%	2	8%
Yes	21	91%	22	92%
Are you currently mentoring another French teacher?				
No	16	70%	17	71%
Yes	7	30%	7	29%
How many years of experience do you have as a French teacher in your state?				
3 years or less	1	4%	1	4%
4 - 7 years	4	17%	5	21%
8 - 11 years	7	30%	4	17%
12 - 15 years	3	13%	2	8%
16 years or more	8	35%	11	46%
For which education level are you currently teaching French?				
Elementary (K - 5 or K - 6)	2	9%	0	0%
Middle School (6 - 8 or 7 - 9)	1	4%	1	4%
High School (9 - 12 or 10 - 12)	11	48%	18	75%
Middle/High School	2	9%	0	0%
All Grades (K - 12)	0	0%	1	4%
Higher Education	6	26%	4	17%
Other	1	4%	0	0%
School Setting				
Urban	10	43%	9	38%
Suburban	6	26%	9	38%
Rural	7	30%	6	25%

Table 8a Cut score Summary by Round of Judgments — French Panel 1

	Round 1	Round 2
Section I: Listening (Max. Raw Score = 30)		
Average	17.58	17.24
SD	2.24	1.90
SEJ	0.47	0.40
Highest	22.05	21.45
Lowest	13.09	14.20
Section II: Reading (Max. Raw Score = 31)		
Average	21.48	21.47
SD	2.86	2.39
SEJ	0.60	0.50
Highest	28.75	27.65
Lowest	15.00	16.20
Section III: Writing (Max. Raw Score = 18)		
Average	9.78	10.26
SD	1.31	1.14
SEJ	0.27	0.24
Highest	12.00	12.00
Lowest	8.00	8.00
Section IV: Speaking (Max. Raw Score = 18)		
Average	9.52	9.57
SD	2.35	1.95
SEJ	0.49	0.41
Highest	16.00	16.00
Lowest	6.00	7.00
Total (Max. Raw Score = 97)		
Average	58.37	58.54
SD	5.33	4.56
SEJ	1.11	0.95
Highest	66.05	65.55
Lowest	45.00	48.20

Table 8b Panelists Cut scores by Round of Judgments — French Panel 1

Panelist	Section I		Section II		Section III		Section IV		Total	
	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2
1	19.90	19.50	23.95	23.95	9.00	11.00	11.00	11.00	63.85	65.45
2	14.90	14.80	21.60	20.60	10.00	11.00	10.00	10.00	56.50	56.40
3	19.00	18.00	20.95	20.95	11.00	11.00	11.00	11.00	61.95	60.95
4	16.25	15.55	23.25	22.70	8.00	8.00	6.00	8.00	53.50	54.25
5	17.00	16.60	19.65	20.25	9.00	9.00	10.00	10.00	55.65	55.85
6	18.60	18.00	22.30	21.80	10.00	11.00	9.00	9.00	59.90	59.80
7	15.20	15.80	17.50	19.20	9.00	12.00	11.00	10.00	52.70	57.00
8	15.00	16.00	15.00	16.20	8.00	9.00	7.00	7.00	45.00	48.20
9	15.85	14.75	20.00	20.05	10.00	9.00	10.00	7.00	55.85	50.80
10	22.05	21.45	25.35	25.00	10.00	10.00	7.00	7.00	64.40	63.45
11	19.60	17.40	21.85	20.95	12.00	12.00	12.00	11.00	65.45	61.35
12	16.20	16.30	19.90	20.10	10.00	10.00	8.00	9.00	54.10	55.40
13	14.65	15.65	18.90	18.50	10.00	10.00	11.00	10.00	54.55	54.15
14	19.25	18.25	23.90	23.60	8.00	8.00	10.00	8.00	61.15	57.85
15	19.20	17.50	23.60	22.55	8.00	10.00	6.00	9.00	56.80	59.05
16	19.45	18.85	22.70	22.40	8.00	10.00	7.00	9.00	57.15	60.25
17	16.20	16.00	19.20	20.10	9.00	10.00	7.00	9.00	51.40	55.10
18	17.30	17.90	28.75	27.65	12.00	12.00	8.00	8.00	66.05	65.55
19	18.90	18.10	20.50	20.60	10.00	10.00	16.00	16.00	65.40	64.70
20	13.90	14.20	20.50	21.80	12.00	11.00	11.00	10.00	57.40	57.00
21	19.55	18.75	21.15	21.35	11.00	11.00	10.00	10.00	61.70	61.10
22	15.80	16.20	19.60	19.60	11.00	11.00	12.00	12.00	58.40	58.80
23	20.60	21.00	24.00	24.00	10.00	10.00	9.00	9.00	63.60	64.00

Table 9a Cut score Summary by Round of Judgments — French Panel 2

	Round 1	Round 2
Section I: Listening (Max. Raw Score = 30)		
Average	18.34	18.10
SD	2.46	1.96
SEJ	0.50	0.40
Highest	22.50	21.70
Lowest	14.60	15.20
Section II: Reading (Max. Raw Score = 31)		
Average	22.80	23.08
SD	2.64	2.29
SEJ	0.54	0.47
Highest	27.40	27.30
Lowest	16.40	17.40
Section III: Writing (Max. Raw Score = 18)		
Average	12.00	12.67
SD	1.35	1.05
SEJ	0.28	0.21
Highest	14.00	15.00
Lowest	9.00	11.00
Section IV: Speaking (Max. Raw Score = 18)		
Average	11.54	12.00
SD	1.61	1.10
SEJ	0.33	0.23
Highest	14.00	14.00
Lowest	8.00	9.00
Total (Max. Raw Score = 97)		
Average	64.68	65.84
SD	6.03	4.68
SEJ	1.23	0.96
Highest	74.25	73.55
Lowest	52.00	56.40

Table 9b Panelists Cut scores by Round of Judgments — French Panel 2

Panelist	Section I		Section II		Section III		Section IV		Total	
	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2
1	16.40	15.90	23.20	23.40	13.00	13.00	13.00	13.00	65.60	65.30
2	21.00	20.60	24.65	23.85	13.00	13.00	12.00	12.00	70.65	69.45
3	16.40	15.20	22.35	22.45	11.00	11.00	8.00	12.00	57.75	60.65
4	15.50	16.30	21.40	22.30	12.00	12.00	12.00	12.00	60.90	62.60
5	22.15	20.55	27.40	27.30	12.00	12.00	12.00	12.00	73.55	71.85
6	15.95	16.15	23.50	24.10	9.00	11.00	9.00	10.00	57.45	61.25
7	15.30	15.50	18.70	19.90	11.00	11.00	8.00	10.00	53.00	56.40
8	17.35	18.25	23.00	23.90	13.00	13.00	10.00	12.00	63.35	67.15
9	19.60	19.70	24.75	25.95	11.00	13.00	11.00	12.00	66.35	70.65
10	16.15	17.95	19.00	21.70	13.00	15.00	11.00	12.00	59.15	66.65
11	19.40	18.40	20.10	19.50	12.00	12.00	10.00	12.00	61.50	61.90
12	17.20	17.85	24.55	24.35	10.00	13.00	11.00	12.00	62.75	67.20
13	20.75	19.95	23.30	22.90	13.00	12.00	13.00	13.00	70.05	67.85
14	19.65	19.80	22.00	22.50	13.00	13.00	12.00	12.00	66.65	67.30
15	20.55	20.15	22.70	22.70	14.00	14.00	12.00	12.00	69.25	68.85
16	15.30	15.85	23.10	23.10	13.00	13.00	13.00	12.00	64.40	63.95
17	19.75	18.05	22.90	23.00	11.00	11.00	11.00	9.00	64.65	61.05
18	14.60	15.30	16.40	17.40	10.00	13.00	11.00	11.00	52.00	56.70
19	20.75	20.20	25.30	25.40	12.00	13.00	13.00	13.00	71.05	71.60
20	18.95	18.45	22.75	22.55	13.00	13.00	12.00	13.00	66.70	67.00
21	15.20	15.80	21.70	21.70	13.00	13.00	13.00	12.00	62.90	62.50
22	18.90	18.10	20.70	21.30	12.00	14.00	14.00	14.00	65.60	67.40
23	20.80	18.65	26.45	25.75	14.00	14.00	13.00	13.00	74.25	71.40
24	22.50	21.70	27.30	26.85	10.00	12.00	13.00	13.00	72.80	73.55

Table 10a Cut scores within 1 and 2 SEMs of the Recommended Cut score — French Panel 1

Recommended Cut score (SEM)		Scale Score Equivalent
59 (4.65)		157
- 2 SEMs	50	145
-1 SEM	54	150
+1 SEM	64	163
+ 2 SEMs	68	169

Table 10b Cut scores within 1 and 2 SEMs of the Recommended Cut score — French Panel 2

Recommended Cut score (SEM)		Scale Score Equivalent
66 (4.54)		166
- 2 SEMs	57	154
-1 SEM	62	161
+1 SEM	71	172
+ 2 SEMs	75	178

Table 10c Cut scores within 1 and 2 SEMs of the Recommended Cut score — Combined French Panels

Recommended Cut score (SEM)		Scale Score Equivalent
63 (4.61)		162
- 2 SEMs	53	149
-1 SEM	58	156
+1 SEM	67	167
+ 2 SEMs	72	174

Note: Consistent with the recommended cut score, the cut scores at the different SEMs have been rounded to the next highest whole number.

Table 11 Specification Judgments — French (Panels 1 & 2 Judgments Combined)

	Very Important		Important		Slightly Important		Not Important	
	N	%	N	%	N	%	N	%
Language, Linguistics, and Comparison								
A. Demonstrating Language Proficiency	35	74%	12	26%	0	0%	0	0
Subtopic 1	22	47%	23	49%	2	4%	0	0
Subtopic 2	36	77%	11	23%	0	0%	0	0
Subtopic 3	27	57%	19	40%	1	2%	0	0%
Subtopic 4	22	47%	24	51%	1	2%	0	0
Subtopic 5	31	66%	16	34%	0	0%	0	0
Subtopic 6	33	70%	12	26%	2	4%	0	0
Subtopic 7	6	13%	37	79%	4	9%	0	0
Subtopic 8	11	23%	32	68%	4	9%	0	0
Subtopic 9	41	87%	6	13%	0	0%	0	0
Subtopic 10	25	53%	18	38%	3	6%	0	0%
Subtopic 11	19	40%	26	55%	2	4%	0	0
B. Understanding Linguistics	21	46%	23	50%	2	4%	0	0
Subtopic 1	18	40%	21	47%	6	13%	0	0
Subtopic 2	24	51%	22	47%	1	2%	0	0
Subtopic 3	24	51%	21	45%	2	4%	0	0
Subtopic 4	23	50%	19	41%	4	9%	0	0
Subtopic 5	24	51%	21	45%	2	4%	0	0
Subtopic 6	27	57%	18	38%	2	4%	0	0
C. Comparison of Target Language with English	19	42%	20	44%	5	11%	0	0%
Subtopic 1	24	51%	17	36%	5	11%	0	0%
Subtopic 2	15	32%	25	53%	6	13%	0	0%
Cultures, Literature, Cross-disciplinary Concepts								
A. Demonstrating Language Proficiency	25	53%	22	47%	0	0%	0	0
Subtopic 1	23	49%	20	43%	4	9%	0	0
Subtopic 2	16	34%	28	60%	3	6%	0	0

Table 12a Final Evaluation — French Panel 1

	Strongly Agree		Agree		Disagree		Strongly Disagree	
	N	%	N	%	N	%	N	%
I understood the purpose of this study.	21	91%	2	9%	0	0%	0	0%
The instructions and explanations provided by the facilitator were clear.	18	78%	5	22%	0	0%	0	0%
The training in the standard setting methods was adequate to give me the information I needed to complete my assignment.	18	78%	5	22%	0	0%	0	0%
The explanation of how the recommended cut scores are computed was clear.	21	91%	2	9%	0	0%	0	0%
The opportunity for feedback and discussion between rounds was helpful.	15	65%	6	26%	2	9%	0	0%
The process of making the standard setting judgments was easy to follow.	15	65%	8	35%	0	0%	0	0%
	Very Influential		Somewhat Influential		Not Influential			
How influential was each of the following factors in guiding your standard setting judgments?	N	%	N	%	N	%		
The definition of the Just Qualified Candidate	20	87%	2	9%	1	4%		
The between-round discussions	10	43%	12	52%	1	4%		
The cut scores of other panel members	19	83%	4	17%	0	0%		
My own professional experience	2	9%	18	78%	3	13%		
	Very Comfortable		Somewhat Comfortable		Somewhat Uncomfortable		Very Uncomfortable	
	N	%	N	%	N	%	N	%
Overall, how comfortable are you with the panel's recommended cut score?	18	78%	5	22%	0	0%	0	0%
	Too Low		About Right		Too High			
Overall, the panel's recommended cut score is:	1	4%	22	96%	0	0%		

Table 12b Final Evaluation — French Panel 2

	Strongly Agree		Agree		Disagree		Strongly Disagree	
	N	%	N	%	N	%	N	%
I understood the purpose of this study.	23	96%	1	4%	0	0%	0	0%
The instructions and explanations provided by the facilitator were clear.	23	96%	1	4%	0	0%	0	0%
The training in the standard setting methods was adequate to give me the information I needed to complete my assignment.	21	88%	3	13%	0	0%	0	0%
The explanation of how the recommended cut scores are computed was clear.	19	79%	5	21%	0	0%	0	0%
The opportunity for feedback and discussion between rounds was helpful.	22	92%	2	8%	0	0%	0	0%
The process of making the standard setting judgments was easy to follow.	21	88%	3	13%	0	0%	0	0%
	Very Influential		Somewhat Influential		Not Influential			
How influential was each of the following factors in guiding your standard setting judgments?	N	%	N	%	N	%		
The definition of the Just Qualified Candidate	19	79%	5	21%	0	0%		
The between-round discussions	15	63%	9	38%	0	0%		
The cut scores of other panel members	3	100%	0	0%	0	0%		
My own professional experience	2	8%	16	67%	6	25%		
	Very Comfortable		Somewhat Comfortable		Somewhat Uncomfortable		Very Uncomfortable	
	N	%	N	%	N	%	N	%
Overall, how comfortable are you with the panel's recommended cut score?	19	79%	4	17%	1	4%	0	0%
	Too Low		About Right		Too High			
Overall, the panel's recommended cut score is:	2	8%	22	92%	0	0%		

APPENDIX F
Results for Praxis World Languages: Spanish

Table 13 Committee Member Demographics — Spanish

	Panel 1		Panel 2	
	N	Percent	N	Percent
Group you are representing				
Teachers	12	67%	12	57%
Administrator/Department Head	2	11%	5	24%
College Faculty	4	22%	4	19%
Other	0	0%	0	0%
Race				
African American or Black	3	17%	4	19%
Alaskan Native or American Indian	0	0%	0	0%
Asian or Asian American	1	6%	0	0%
Native Hawaiian or Other Pacific Islander	0	0%	1	5%
White	9	50%	8	38%
Hispanic	5	28%	8	38%
Gender				
Female	12	67%	16	76%
Male	6	33%	5	24%
In which language are you most fluent?				
English	13	72%	10	48%
Spanish	1	6%	2	10%
English and Spanish about the same	4	22%	9	43%
Other	0	0%	0	0%
Are you certified as a Spanish teacher in your state?				
No	4	22%	5	24%
Yes	14	78%	16	76%
Are you currently teaching Spanish in your state?				
No	1	6%	4	19%
Yes	17	94%	17	81%
Are you currently mentoring another Spanish teacher?				
No	11	61%	10	48%
Yes	7	39%	11	52%
How many years of experience do you have as a Spanish teacher in your state?				
3 years or less	0	0%	0	0%
4 - 7 years	3	17%	1	5%
8 - 11 years	4	22%	8	38%
12 - 15 years	2	11%	2	10%
16 years or more	9	50%	10	48%
For which education level are you currently teaching Spanish?				
Elementary (K - 5 or K - 6)	0	0%	0	0%
Middle School (6 - 8 or 7 - 9)	0	0%	1	5%
High School (9 - 12 or 10 - 12)	10	56%	12	57%
Middle/High School	2	11%	0	0%
All Grades (K - 12)	1	6%	2	10%
Higher Education	5	28%	6	29%
School Setting				
Urban	9	50%	10	48%
Suburban	2	11%	7	33%
Rural	7	39%	4	19%

Table 14a Cut score Summary by Round of Judgments — Spanish Panel 1

	Round 1	Round 2
Section I: Listening (Max. Raw Score = 30)		
Average	20.18	20.23
SD	2.43	2.34
SEJ	0.57	0.55
Highest	23.95	24.05
Lowest	14.05	14.05
Section II: Reading (Max. Raw Score = 31)		
Average	22.15	22.21
SD	2.56	2.63
SEJ	0.60	0.62
Highest	25.20	25.20
Lowest	15.25	14.75
Section III: Writing (Max. Raw Score = 18)		
Average	11.50	11.78
SD	1.42	1.40
SEJ	0.33	0.33
Highest	14.00	14.00
Lowest	9.00	9.00
Section IV: Speaking (Max. Raw Score = 18)		
Average	11.44	11.33
SD	1.50	1.24
SEJ	0.35	0.29
Highest	15.00	13.00
Lowest	9.00	9.00
Total (Max. Raw Score = 97)		
Average	62.27	65.54
SD	5.94	5.99
SEJ	1.40	1.41
Highest	77.65	76.25
Lowest	51.30	49.80

Table 14b Panelists Cut scores by Round of Judgments — Spanish Panel 1

Panelist	Section I		Section II		Section III		Section IV		Total	
	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2
1	19.80	19.80	19.30	19.30	10.00	10.00	11.00	12.00	60.10	61.10
2	21.20	21.50	22.85	22.85	12.00	12.00	11.00	11.00	67.05	67.35
3	20.25	20.35	23.75	23.75	9.00	9.00	11.00	11.00	64.00	64.10
4	18.40	18.70	20.85	20.85	12.00	13.00	9.00	9.00	60.25	61.55
5	19.35	19.95	25.05	25.05	10.00	12.00	11.00	12.00	65.40	69.00
6	23.45	24.05	25.20	25.20	14.00	14.00	15.00	13.00	77.65	76.25
7	17.65	17.65	21.60	21.60	12.00	12.00	12.00	12.00	63.25	63.25
8	14.05	14.05	15.25	14.75	11.00	11.00	11.00	10.00	51.30	49.80
9	23.50	23.60	22.95	22.95	13.00	13.00	13.00	13.00	72.45	72.55
10	21.55	20.85	24.05	23.55	10.00	10.00	10.00	11.00	65.60	65.40
11	22.75	22.35	23.85	23.75	12.00	13.00	12.00	12.00	70.60	71.10
12	19.65	20.55	25.05	25.00	14.00	14.00	13.00	13.00	71.70	72.55
13	19.40	19.60	20.50	20.20	10.00	11.00	9.00	9.00	58.90	59.80
14	21.20	19.80	21.15	20.65	12.00	12.00	12.00	12.00	66.35	64.45
15	19.80	19.80	23.70	24.20	10.00	10.00	11.00	11.00	64.50	65.00
16	18.50	18.70	21.30	21.80	12.00	12.00	13.00	12.00	64.80	64.50
17	18.80	19.45	19.05	20.05	12.00	12.00	12.00	11.00	61.85	62.50
18	23.95	23.35	23.20	24.20	12.00	12.00	10.00	10.00	69.15	69.55

Table 15a Cut score Summary by Round of Judgments — Spanish Panel 2

	Round 1	Round 2
Section I: Listening (Max. Raw Score = 30)		
Average	21.76	21.47
SD	2.63	2.19
SEJ	0.57	0.48
Highest	27.00	25.45
Lowest	16.30	17.40
Section II: Reading (Max. Raw Score = 31)		
Average	22.90	22.89
SD	3.27	2.74
SEJ	0.71	0.60
Highest	28.45	26.40
Lowest	15.10	16.20
Section III: Writing (Max. Raw Score = 18)		
Average	12.19	12.19
SD	1.17	1.29
SEJ	0.25	0.28
Highest	14.00	15.00
Lowest	10.00	10.00
Section IV: Speaking (Max. Raw Score = 18)		
Average	11.48	11.48
SD	2.42	2.23
SEJ	0.53	0.49
Highest	14.00	15.00
Lowest	6.00	6.00
Total (Max. Raw Score = 97)		
Average	68.32	68.02
SD	5.97	5.91
SEJ	1.30	1.29
Highest	76.65	80.50
Lowest	51.40	54.60

Table 15b Panelists Cut scores by Round of Judgments — Spanish Panel 2

Panelist	Section I		Section II		Section III		Section IV		Total	
	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2	Rd 1	Rd 2
1	23.00	22.85	17.35	22.75	13.00	13.00	10.00	11.00	63.35	69.60
2	23.70	22.20	23.60	23.05	12.00	12.00	12.00	11.00	71.30	68.25
3	22.50	21.80	22.30	22.10	12.00	12.00	12.00	12.00	68.80	67.90
4	22.80	22.85	23.45	23.65	13.00	13.00	13.00	13.00	72.25	72.50
5	27.00	25.45	28.45	26.40	12.00	12.00	6.00	7.00	73.45	70.85
6	17.85	17.70	20.15	19.65	11.00	11.00	10.00	10.00	59.00	58.35
7	21.65	20.90	22.40	21.90	13.00	13.00	10.00	10.00	67.05	65.80
8	22.75	22.10	24.80	24.55	11.00	11.00	13.00	13.00	71.55	70.65
9	21.55	20.85	26.15	25.50	10.00	10.00	10.00	10.00	67.70	66.35
10	22.85	20.85	25.25	24.35	13.00	13.00	9.00	10.00	70.10	68.20
11	23.30	22.45	25.20	24.60	11.00	11.00	6.00	6.00	65.50	64.05
12	23.45	23.35	24.95	24.75	13.00	13.00	14.00	14.00	75.40	75.10
13	22.55	22.40	19.45	19.15	12.00	12.00	13.00	13.00	67.00	66.55
14	17.80	19.85	25.30	25.35	14.00	14.00	14.00	14.00	71.10	73.20
15	16.30	17.40	15.10	16.20	10.00	11.00	10.00	10.00	51.40	54.60
16	22.80	22.85	22.70	22.60	13.00	13.00	13.00	13.00	71.50	71.45
17	20.35	19.55	22.30	21.35	12.00	10.00	13.00	12.00	67.65	62.90
18	23.80	23.45	26.40	26.40	13.00	13.00	13.00	12.00	76.20	74.85
19	23.30	24.45	25.35	26.05	14.00	15.00	14.00	15.00	76.65	80.50
20	16.75	17.55	20.95	21.05	13.00	13.00	14.00	13.00	64.70	64.60
21	20.95	20.05	19.20	19.20	11.00	11.00	12.00	12.00	63.15	62.25

Table 16a Cut scores within 1 and 2 SEMs of the Recommended Cut score — Spanish Panel 1

Recommended Cut score (SEM)		Scale Score Equivalent
66 (4.44)		167
- 2 SEMs	57	155
-1 SEM	62	162
+1 SEM	70	172
+ 2 SEMs	75	179

Table 16b Cut scores within 1 and 2 SEMs of the Recommended Cut score — Spanish Panel 2

Recommended Cut score (SEM)		Scale Score Equivalent
69 (4.33)		171
- 2 SEMs	60	159
-1 SEM	64	164
+1 SEM	73	176
+ 2 SEMs	77	181

Table 16c Cut scores within 1 and 2 SEMs of the Recommended Cut score — Combined Spanish Panels

Recommended Cut score (SEM)		Scale Score Equivalent
67 (4.38)		168
- 2 SEMs	58	156
-1 SEM	63	163
+1 SEM	72	175
+ 2 SEMs	76	180

Note: Consistent with the recommended cut score, the cut scores at the different SEMs have been rounded to the next highest whole number.

Table 17 Specification Judgments — Spanish (Panels 1 & 2 Judgments Combined)

	Very Important		Important		Slightly Important		Not Important	
	N	%	N	%	N	%	N	%
Language, Linguistics, and Comparison								
A. Demonstrating Language Proficiency	31	84%	6	16%	0	0%	0	0%
Subtopic 1	26	67%	13	33%	0	0%	0	0%
Subtopic 2	29	74%	10	26%	0	0%	0	0%
Subtopic 3	19	49%	19	49%	1	3%	0	0%
Subtopic 4	20	51%	19	49%	0	0%	0	0%
Subtopic 5	24	62%	15	38%	0	0%	0	0%
Subtopic 6	26	67%	12	31%	1	3%	0	0%
Subtopic 7	10	26%	28	72%	1	3%	0	0%
Subtopic 8	12	31%	27	69%	0	0%	0	0%
Subtopic 9	33	85%	5	13%	1	3%	0	0%
Subtopic 10	18	46%	21	54%	0	0%	0	0%
Subtopic 11	14	36%	21	54%	4	10%	0	0%
B. Understanding Linguistics	20	53%	17	45%	0	0%	0	0%
Subtopic 1	20	51%	18	46%	1	3%	0	0%
Subtopic 2	16	41%	20	51%	3	8%	0	0%
Subtopic 3	17	44%	22	56%	0	0%	0	0%
Subtopic 4	23	59%	13	33%	2	5%	0	0%
Subtopic 5	30	77%	8	21%	1	3%	0	0%
Subtopic 6	19	49%	18	46%	2	5%	0	0%
C. Comparison of Target Language with English	13	35%	22	59%	2	5%	0	0%
Subtopic 1	15	38%	21	54%	3	8%	0	0%
Subtopic 2	11	28%	24	62%	4	10%	0	0%
Cultures, Literature, Cross-disciplinary Concepts								
A. Demonstrating Language Proficiency	15	41%	21	57%	1	3%	0	0%
Subtopic 1	15	39%	20	53%	2	5%	0	0%
Subtopic 2	17	45%	18	47%	3	8%	0	0%

Table 18a Final Evaluation — Spanish Panel 1

	Strongly Agree		Agree		Disagree		Strongly Disagree	
	N	%	N	%	N	%	N	%
I understood the purpose of this study.	18	100%	0	0%	0	0%	0	0%
The instructions and explanations provided by the facilitator were clear.	16	89%	2	11%	0	0%	0	0%
The training in the standard setting methods was adequate to give me the information I needed to complete my assignment.	15	83%	3	17%	0	0%	0	0%
The explanation of how the recommended cut scores are computed was clear.	13	72%	5	28%	0	0%	0	0%
The opportunity for feedback and discussion between rounds was helpful.	16	94%	1	6%	0	0%	0	0%
The process of making the standard setting judgments was easy to follow.	12	67%	6	33%	0	0%	0	0%
	Very Influential		Somewhat Influential		Not Influential			
How influential was each of the following factors in guiding your standard setting judgments?	N	%	N	%	N	%		
The definition of the Just Qualified Candidate	18	100%	0	0%	0	0%		
The between-round discussions	13	72%	5	28%	0	0%		
The cut scores of other panel members	6	35%	9	53%	2	12%		
My own professional experience	16	89%	2	11%	0	0%		
	Very Comfortable		Somewhat Comfortable		Somewhat Uncomfortable		Very Uncomfortable	
	N	%	N	%	N	%	N	%
Overall, how comfortable are you with the panel's recommended cut score?	11	61%	5	28%	2	11%	0	0%
	Too Low		About Right		Too High			
Overall, the panel's recommended cut score is:	1	6%	15	83%	2	11%		

Table 18b Final Evaluation — Spanish Panel 2

	Strongly Agree		Agree		Disagree		Strongly Disagree	
	N	%	N	%	N	%	N	%
I understood the purpose of this study.	18	90%	2	10%	0	0%	0	0%
The instructions and explanations provided by the facilitator were clear.	19	95%	1	5%	0	0%	0	0%
The training in the standard setting methods was adequate to give me the information I needed to complete my assignment.	19	95%	1	5%	0	0%	0	0%
The explanation of how the recommended cut scores are computed was clear.	18	90%	2	10%	0	0%	0	0%
The opportunity for feedback and discussion between rounds was helpful.	20	100%	0	0%	0	0%	0	0%
The process of making the standard setting judgments was easy to follow.	16	80%	4	20%	0	0%	0	0%
	Very Influential		Somewhat Influential		Not Influential			
How influential was each of the following factors in guiding your standard setting judgments?	N	%	N	%	N	%		
The definition of the Just Qualified Candidate	18	90%	2	10%	0	0%		
The between-round discussions	11	55%	7	35%	2	10%		
The cut scores of other panel members	0	0%	18	90%	2	10%		
My own professional experience	16	80%	3	15%	1	5%		
	Very Comfortable		Somewhat Comfortable		Somewhat Uncomfortable		Very Uncomfortable	
	N	%	N	%	N	%	N	%
Overall, how comfortable are you with the panel's recommended cut score? ¹¹								
	Too Low		About Right		Too High			
Overall, the panel's recommended cut score is: ¹¹								

¹¹ Due to technical problems during the study, panelists were not able to review and judge their comfort level with the overall cut score following Round 2.

APPENDIX G
Praxis World Languages Content Specifications

Language, Linguistics, and Comparisons

Demonstrating Language Proficiency

1. Knows how to communicate in the target language with native speakers unaccustomed to dealing with nonnative speakers, with sufficient accuracy, clarity, and precision to convey the intended message
2. Knows how to communicate in the interpersonal mode (speaking) by participating actively in informal and formal conversations on topics covering home, school, leisure activities, and current events
3. Knows how to communicate in the interpersonal mode (writing) in written exchanges on daily topics
4. Comprehends in the interpretive mode (listening) main ideas and supporting details of audio segments such as news items, short stories, social notices, and reports on familiar topics that deal with factual information
5. Comprehends in the interpretive mode (reading) main ideas and supporting details of printed texts such as news items, short stories, social notices, and reports on familiar topics that deal with factual information
6. Knows how to negotiate meaning in order to sustain an interaction
7. Knows how to move beyond literal comprehension in the interpretive mode (listening) by inferring the meaning of unfamiliar words and phrases in new contexts, inferring and interpreting the author's intent, and offering a personal interpretation of the message
8. Knows how to move beyond literal comprehension in the interpretive mode (reading) by inferring the meaning of unfamiliar words and phrases in new contexts, inferring and interpreting the author's intent, and offering a personal interpretation of the message
9. Understands the gist of normal conversational speech on a variety of topics
10. Knows how to communicate in the presentational mode (writing) by writing routine social correspondence, as well as coherent narratives, descriptions, and summaries about familiar topics of a factual nature in paragraph length in present, past, and future time
11. Knows how to communicate orally in the presentational mode (speaking) by delivering oral presentations on familiar literary or cultural topics and incorporating extra linguistic support to facilitate oral presentations that are extemporaneous or prepared but not read

Understanding Linguistics

1. Understands the rules of the sound system of the target language (i.e., recognizing phonemes and allophones)
2. Recognizes key cohesive devices (conjunctions and adverbs) used in connected discourse
3. Understands high-frequency idiomatic expressions and can infer meaning of words and sentences
4. Knows how to explain the rules that govern the formation of words and sentences in the target language
5. Knows how to exemplify the rules with examples from the target languages, such as the verbal system, pronouns, agreement, word order, interrogatives, both in terms of regularities and irregularities
6. Knows how to identify and use the pragmatic and sociolinguistics conventions and register (formal and informal forms of address)

Comparison of Target Language with English

1. Knows how to identify similarities and differences between the target language and English
2. Knows how to contrast syntactical patterns of simple sentences and questions with those of English

Cultures, Literatures, Cross-Disciplinary Concepts

Demonstrating Cultural Understandings

1. Knows the three Ps:
 - Perspectives (such as attitudes, ideas, and values)
 - Practices (patterns of behavior and social interaction, such as greetings, turn taking, and rites of passage) and
 - Products (such as tools, foods, law, and music)
2. Recognizes the value and role of authentic literary and cultural texts—such as songs, poems, rhymes and chants, children's books, narrative text, and novels—and usage of those texts to interpret and reflect on the perspectives of the target cultures